

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-305790

(43)Date of publication of application : 05.11.1999

(51)Int.Cl.

G10L 3/00
 G10L 3/00
 G10L 3/00
 // B60R 16/02

(21)Application number : 10-113393

(71)Applicant : DENSO CORP

(22)Date of filing : 23.04.1998

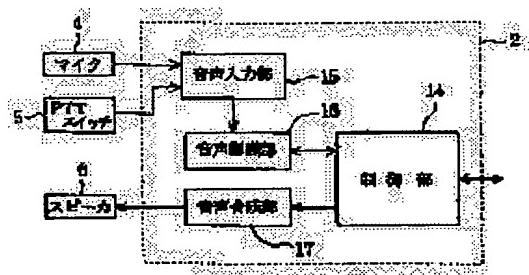
(72)Inventor : KAWAKAMI HIDEHIKO

(54) VOICE RECOGNITION DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To improve a recognition rate in voice recognition by constituting a category of an input voice specifiable and performing the voice recognition based on a vocabulary belonging to the specified category in the vocabulary in a recognition dictionary and the input voice.

SOLUTION: A voice input part 15 inputs a voice uttered from a user through a microphone 4 to output the voice data to a voice recognition part 16. At this time, the voice input part 15 outputs the voice data to the voice recognition part 16 only while the user is depression operating a PTT switch 5. The voice recognition part 16 voice recognition processes the voice data (inputted voice) imparted from the voice input part 15 according to an instruction from a control part 14 to output the voice recognition result to the control part 14. A collation part collates (recognizes) for the voice data imparted from the voice input part 15 by using the dictionary data stored in a dictionary part. Then, the collation part outputs a comparison objective pattern candidate (recognition objective vocabulary) with the highest calculated similarity to the control part 14 as the recognition result.



LEGAL STATUS

[Date of request for examination] 30.05.2000

[Date of sending the examiner's decision of rejection] 27.05.2003

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-305790

(43) 公開日 平成11年(1999)11月5日

(51) Int.Cl. ⁸ G 1 0 L 3/00	識別記号 5 6 1 5 5 1 5 7 1 // B 6 0 R 16/02	F I G 1 0 L 3/00 5 6 1 C 5 6 1 G 5 5 1 J 5 7 1 C B 6 0 R 16/02 審査請求 未請求 請求項の数3 O L (全 7 頁)
---	---	---

(21) 出願番号 特願平10-113393

(22) 出願日 平成10年(1998)4月23日

(71) 出願人 000004260

株式会社デンソー

愛知県刈谷市昭和町1丁目1番地

(72) 発明者 川上 英彦

愛知県刈谷市昭和町1丁目1番地 株式会
社デンソー内

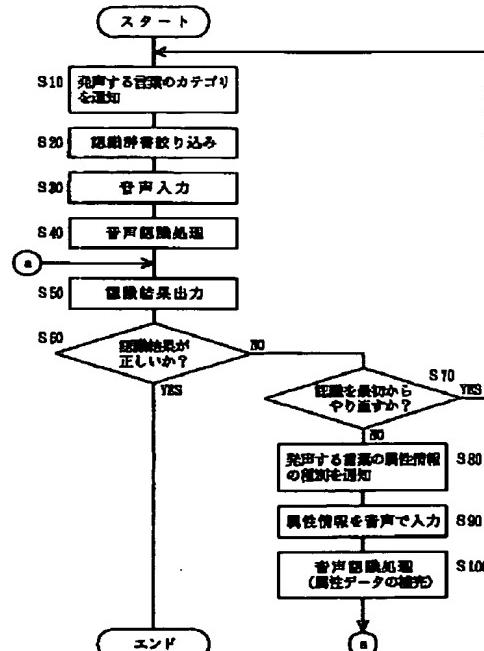
(74) 代理人 弁理士 佐藤 強

(54) 【発明の名称】 音声認識装置

(57) 【要約】

【課題】 音声認識の認識率を向上させて、ユーザーが不快に感ずることを極力防止する。

【解決手段】 本発明の音声認識装置は、入力された音声と認識辞書に基づいて音声認識を実行するように構成されたものにおいて、入力音声で表わされる語彙のカテゴリを指定可能に構成すると共に、認識辞書内の語彙のうちで指定されたカテゴリに属する語彙と入力された音声に基づいて音声認識を行うように構成したものである。この構成によれば、音声認識に用いられる認識辞書の語彙の個数が絞られると共に、同じカテゴリの語彙となるので、音声認識の認識率が向上する。



【特許請求の範囲】

【請求項1】 入力された音声と認識辞書とに基づいて音声認識を実行するように構成された音声認識装置において、

入力音声で表わされる語彙のカテゴリを指定するカテゴリ指定手段と、前記認識辞書内の語彙のうちで前記指定されたカテゴリに属する語彙と入力音声とに基づいて音声認識を行う音声認識手段とを備えたことを特徴とする音声認識装置。

【請求項2】 入力された音声と認識辞書とに基づいて音声認識を実行するように構成された音声認識装置において、

前記認識辞書内の語彙を、使用頻度によって複数の語彙グループに分けると共に、これら複数の語彙グループに使用頻度が高い順に優先順位を付けるように構成し、そして、

入力された音声の認識候補として複数の語彙が output されたときに、前記優先順位が高い語彙グループに属する認識候補を認識結果として出力する音声認識手段を備えたことを特徴とする音声認識装置。

【請求項3】 入力された音声と認識辞書とに基づいて音声認識を実行するように構成された音声認識装置において、

入力音声で表わされる語彙の属性情報を入力する入力手段と、

前記入力された属性情報を参照して音声認識を実行する音声認識手段とを備えたことを特徴とする音声認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、入力された音声と認識辞書とに基づいて音声認識を実行するように構成された音声認識装置に関する。

【0002】

【従来の技術】この種の音声認識装置においては、ユーザーが発した音声を入力し、この入力した音声と、認識辞書に記憶されている複数の比較対象パターン候補とを比較（照合）して、ユーザー発生音声と比較対象パターン候補の一一致度合い、即ち、類似度を計算するように構成されている。そして、これら計算した類似度の中で最も高い類似度を持つ比較対象パターン候補を、認識結果として出力するように構成されている。

【0003】このような構成の音声認識装置をナビゲーションシステムに組み込むと、ナビゲーションシステムに目的地等を入力する場合に、音声による入力が可能となる。これにより、ナビゲーションシステムを音声によって操作可能となるので、運転中のユーザーにとってかなり利用し易い便利な装置となる。

【0004】

【発明が解決しようとする課題】しかしながら、上記従

来構成の音声認識装置においては、発音が似ている目的地や名称等を音声認識するような場合、認識結果が誤っていることがときどきある。特に、近年、認識可能な語彙の数が非常に増えているため、誤認識が発生したとき、ユーザーが発した語彙のカテゴリと、誤認識された語彙のカテゴリとが全く異なるようなケースが起こるようになつた。即ち、誤認識された語彙が、ユーザーの予期や意図に大きく反してしまうことがあり、このような場合、ユーザーは音声認識装置に対して不快感や不審感を抱くおそれがあった。

【0005】そこで、本発明の目的は、音声認識の認識率を向上させることができ、ユーザーが不快に感ずることを極力防止できる音声認識装置を提供することにある。

【0006】

【課題を解決するための手段】請求項1の発明によれば、入力音声のカテゴリを指定可能に構成し、そして、認識辞書内の語彙のうちで指定されたカテゴリに属する語彙と入力音声とに基づいて音声認識を行うように構成した。この構成の場合、音声認識に用いられる認識辞書の語彙の個数が絞られると共に、同じカテゴリの語彙となるので、音声認識の認識率が向上する。また、たとえ誤認識が発生したとしても、誤認識された語彙が同じカテゴリの語彙であるから、ユーザーに不信感を与えない。

【0007】請求項2の発明においては、認識辞書内の語彙を、使用頻度によって複数の語彙グループに分けると共に、これら複数の語彙グループに使用頻度が高い順に優先順位を付けるように構成し、そして、入力された音声の認識候補として複数の語彙が output されたときに、優先順位が高い語彙グループに属する認識候補を認識結果として出力するように構成した。この場合、優先順位が高い語彙グループに属する言葉は使用頻度が高いから、それだけ、音声認識の認識率が向上する。

【0008】請求項3の発明では、入力音声で表わされる語彙の属性情報を入力可能に構成し、この入力された属性情報を参照して音声認識を実行するように構成した。この構成によれば、語彙の属性情報を参照して音声認識を実行するため、音声認識の認識率が向上する。

【0009】

【発明の実施の形態】以下、本発明をカーナビゲーションシステムに適用した一実施例について図面を参照しながら説明する。まず、図2はカーナビゲーションシステム1の全体構成を概略的に示すブロック図である。この図2に示すように、カーナビゲーションシステム1は、音声認識装置2とナビゲーション装置3とを備えて構成されている。上記音声認識装置2には、マイク4とP.T.T（Push-To-Talk）スイッチ5とスピーカ6とが接続されている。

【0010】また、ナビゲーション装置3の具体的構成

を、図3に示す。この図3において、ナビゲーション装置3の制御回路7は、マイクロコンピュータを含んで構成されており、ナビゲーション装置3の運転全般を制御する機能を有している。この制御回路7には、位置検出器8、地図データ入力器9、操作スイッチ群10、外部メモリ11、表示装置12及びリモコンセンサ13が接続されている。更に、制御回路7には、上記音声入力装置2(の制御部14(図4参照))が接続されている。

【0011】ここで、位置検出器8は、地磁気センサやジャイロスコープや距離センサやGPS受信機(いずれも図示しない)等を組み合わせたもの、または、その一部で構成されている。この位置検出器8は、本実施例のカーナビゲーションシステム1を搭載した車両の現在位置を検出して現在位置検出信号を出力するように構成されている。また、地図データ入力器9は、地図データやマップマッチングデータ等を入力するための装置である。上記地図データ等のデータは、例えばCD-ROMなどからなる記録媒体に記録されている。

【0012】更に、表示装置12は、例えば液晶ディスプレイ等で構成されており、カラー表示が可能で地図等を明確に表示できるものである。操作スイッチ群10は、上記表示装置12の画面の上面に設けられたタッチスイッチ(タッチパネル)と、上記画面の周辺部に設けられたメカニカルなプッシュスイッチ等から構成されている。リモコンセンサ13は、ユーザーにより操作されるリモコン13aから送信された送信信号を受信する受信機である。そして、制御回路7は、ユーザーが操作スイッチ群10やリモコン13aを操作することにより目的地を設定したときに、現在位置からその目的地までの最適経路を自動的に選択設定する機能や、現在位置を地図上に位置付けるマップマッチング処理を実行する機能等を有している。

【0013】また、上記目的地等を設定する場合やナビゲーション装置3のコマンド等を入力する場合に、ユーザーは、操作スイッチ群10やリモコン13aを操作する代わりに、音声認識装置2を用いて音声で入力することが可能に構成されている。以下、上記音声認識装置2について、図4を参照して説明する。

【0014】図4に示すように、音声認識装置2は、制御部14、音声入力部15、音声認識部16及び音声合成部17から構成されている。ここで、制御部14は、音声認識装置2の動作全般を制御する機能を有している。上記制御部14は、上記ナビゲーション装置3の制御回路7に接続されており、これにより、制御回路7との間でデータの授受を行うように構成されている。

【0015】また、音声入力部15は、ユーザーが発した音声をマイク4を介して入力し、音声データ(例えばデジタルデータ)を音声認識部16へ出力するように構成されている。この場合、音声入力部15は、ユーザーがPTTスイッチ5を押し下げ操作している間だけ、音

声データを音声認識部16へ出力するように構成されている。即ち、PTTスイッチ8が操作されている間だけ、ユーザーが発した音声の音声認識処理が実行されるように構成されている。

【0016】そして、音声認識部16は、上記音声入力部15から与えられた音声データ(入力した音声)を制御部14からの指示に従って音声認識処理を行い、その音声認識結果を制御部14へ出力するように構成されている。即ち、音声認識部16が音声認識手段を構成している。上記音声認識部16は、具体的には、図5に示すように、照合部18及び辞書部19から構成されている。上記辞書部19には、認識対象語彙(認識対象の言葉)及びこの認識対象語彙のツリー構造(周知のデータ構造)から構成された辞書データが記憶されている。この辞書データが認識辞書を構成している。

【0017】ここで、辞書データ内に記憶されている認識対象語彙のデータ、即ち、比較対象パターン候補のデータには、都道府県名、市区町村名、位置、種別及びランクのデータ(いわゆる属性データ)が付加されている。この属性データが属性情報を構成している。また、辞書データ内に記憶されている語彙(認識対象語彙)の個数は、10万ないし20万語というレベルの膨大な数である。更に、辞書データ内の認識対象語彙は、例えば3つのカテゴリに分類されている。この3つのカテゴリは、本実施例の場合、コマンド(地図拡大、地図縮小、スクロール等のナビゲーション装置3を操作するための操作コマンド)と、住所と、施設名である。

【0018】また、照合部18は、音声入力部15から与えられた音声データに対して、上記辞書部19に記憶されている辞書データを用いて照合(認識)を行うように構成されている。この場合、まず、音声データと辞書データ内の複数の比較対象パターン候補とを比較して例えば類似度(即ち、両者の一致度合いを計算した値)を計算する。尚、この類似度を計算する処理は、既に知られている照合処理用の制御プログラム(周知のアルゴリズム)を使用して実行されるように構成されている。

【0019】そして、照合部18は、上記計算した類似度が最も高い比較対象パターン候補(認識対象語彙)を、認識結果として制御部14へ出力するように構成されている。ここで、照合部18は、ユーザーによりこれから発声される語彙(言葉)のカテゴリが指定されたときには、辞書データ内に記憶されている認識対象語彙のうちの上記指定されたカテゴリ内に属する語彙と入力音声とを照合するように構成されている。即ち、ユーザーによるカテゴリの指定により、照合に使用する認識対象語彙を絞り込むように構成されている。

【0020】また、ユーザーによるカテゴリの指定は、PTTスイッチ5を設定された回数だけ押圧操作することにより行なわれるよう構成されている。PTTスイッチ5を例えば1回押したときにコマンドの指定にな

り、2回押したときに住所の指定になり、3回押したときに施設名の指定になる。この場合、PTTスイッチ5がカテゴリ指定手段を構成している。尚、上記PTTスイッチ5に代えて、カテゴリ指定キー（例えばコマンドキー、住所キー、施設キー）を操作スイッチ群10の中に設け、各指定キーを操作することによりカテゴリの指定を実行するように構成しても良い。

【0021】尚、制御部14内には、記憶部20が設けられており、この記憶部20にはユーザーの嗜好を反映した認識ルールや、語彙の認識のし易さ及び語彙の認識のし難さ等を表わす認識データ等を記憶させることができになっている。そして、照合部18は、記憶部20内の認識ルールや認識データ等を参照しながら、上記した照合処理を実行するようにも構成されている。

【0022】一方、音声合成部17は、発声させたい音声を表わすデータ（例えば仮名文字等から構成されたテキストデータ）を制御部14から受けけると、この音声データから音声を合成するように構成されている。そして、音声合成部17は、上記合成した音声をスピーカ6から出力して発声させるように構成されている。

【0023】次に、上記構成の作用、具体的には、目的地の設定をユーザーが音声で行う場合の動作について、図1も参照して説明する。図1のフローチャートは、音声認識装置2を動作させる制御プログラムのうちの音声認識処理を実行する制御部分の内容を示している。

【0024】まず、図1のステップS10では、ユーザーによって、これから発声する語彙（言葉）のカテゴリが指定（通知）される。この場合、PTTスイッチ5が1度操作されるとカテゴリとしてコマンドが指定され、2度操作されると住所が指定され、3度操作されると施設が指定される。続いて、ステップS20へ進み、認識辞書（辞書データ）内の語彙の中の上記指定されたカテゴリに属する語彙に基づいて照合（音声認識）を行うよう設定する。即ち、認識辞書を絞り込むように構成されている。

【0025】そして、ユーザーが発声する音声を入力した後（ステップS30）、この入力した音声と上記絞り込んだ認識辞書（辞書データの語彙の中の上記指定されたカテゴリに属する語彙）とに基づいて音声認識を行う（ステップS40）。この場合、指定されたカテゴリによって、音声認識の対象となる辞書データが絞り込まれるから、換言すると、音声認識の対象となる辞書データの語彙数が少なくなるから、認識率が向上するようになる。続いて、音声認識結果を出力する処理、例えば音声認識した語彙を表示装置12に表示したり、音声認識した語彙の音声を合成してスピーカから発声したりした後（ステップS50）、認識結果が正しいか否かをユーザーに問う（ステップS60）。

【0026】ここで、認識結果が正しいという応答がユーザーからあった場合には、ステップS60にてYES

へ進み、音声認識処理を終了する。これに対し、認識結果が正しくないという応答がユーザーからあった場合には、ステップS60にてNOへ進み、音声認識を再び実行する。尚、ステップS60におけるユーザーの応答は、「はい」、「いいえ」という音声で応答しても良いし、操作スイッチ群10に設けられた「YES」キー、「NO」キーを操作して応答しても良い。

【0027】さて、音声認識を再び実行する場合には、音声認識を最初からやり直すか否かを問う（ステップS70）。ここで、音声認識を最初からやり直すという応答がユーザーからあった場合には、ステップS70にてYESへ進み、ステップS10へ戻る。尚、ステップS70におけるユーザーの応答は、「はい」、「いいえ」という音声で応答しても良いし、操作スイッチ群10に設けられた「YES」キー、「NO」キーを操作して応答しても良い。

【0028】一方、音声認識を最初からやり直さないという応答がユーザーからあった場合には、ステップS70にてNOへ進む。そしてこの場合には、ユーザーは、発声する語彙（言葉）の属性情報（属性データ）を指定する。まず、ステップS80へ移行し、属性情報の種別を指定（通知）する。ここでは、ユーザーは、操作スイッチ群10に設けられた属性情報を指定するキー（例えば住所キー、種別キー、ランクキーなど）を操作して指定するように構成されている。尚、属性情報の種別の指定を、音声で入力するように構成しても良い。この場合、例えば「スイッチ住所」、「スイッチ種別」、「スイッチランク」等の語彙を音声で入力できるように構成することが好みしい。

【0029】続いて、ステップS90へ進み、属性情報のデータを音声で入力する。この場合、属性情報の種別が住所であれば、「愛知県」とか、「岐阜県」とか、「名古屋市」というような語彙（データ）を音声認識により入力する。そして、ステップS100へ進み、上記指定された属性情報を参照して、音声認識を再び行う。この場合、ステップS40の音声認識処理で、数個の認識対象候補があったとすれば、その中から認識結果を1つ選ぶ際に、上記指定された属性情報を参照しながら選ぶように構成されている。即ち、指定された属性情報によって、音声認識処理を補完するように構成されており、これにより、認識率が向上するようになっている。そして、ステップS50へ進み、音声認識した結果を出力するように構成されている。

【0030】このような構成の本実施例においては、入力音声のカテゴリを指定可能に構成すると共に、認識辞書内の語彙のうちの指定されたカテゴリに属する語彙と入力された音声とに基づいて音声認識を行うように構成した。このため、音声認識に用いられる認識辞書の語彙の個数が絞られると共に、同じカテゴリの語彙となるので、音声認識の認識率が向上する。特に、誤認識が起

った場合でも、ユーザーが発した語彙のカテゴリと、誤認識された語彙のカテゴリと同じであるから、誤認識された語彙が、ユーザーの予期や意図に大きく反してしまうことがなくなる。従って、音声認識装置に対してユーザーが不快感や不審感をあまり持たなくなる。

【0031】また、上記実施例では、誤認識が発生した後、音声認識を再び行う場合に、入力される音声の語彙の属性情報（都道府県名、市区町村名、位置、種別及びランクのデータ）を入力可能に構成すると共に、この入力された属性情報に基づいて音声認識を実行するように構成した。この構成によれば、語彙の属性情報に基づいて音声認識を実行するため、音声認識の認識率がより一層向上する。

【0032】具体的には、ユーザーが「名古屋球場」を意図して発生した場合に、1回目の音声認識の照合により、3個の認識対象候補が求められ、更に、これらの類似度の点数が次の通り計算されたとする。3個の認識対象候補は、「沖縄県、名護野球場、野球場、90点」、「愛知県、名古屋球場、野球場、89点」、「愛知県、名古屋城、城、88点」であったとする。この場合、名護野球場が認識結果として出力される。このとき、ユーザーは、認識が誤っていることを応答し、再度音声認識を実行するために、属性情報を入力する。この場合、例えば住所キーを操作して、属性情報の種別として住所を指定した後、属性情報のデータとして「愛知県」を音声で入力したとする。すると、この後の音声認識処理により、名古屋球場が認識結果として出力されるようになる。

【0033】尚、上記実施例では、PTTスイッチ5を操作する回数によりカテゴリを指定する構成としたが、これに限られるものではなく、カテゴリを指定する操作スイッチを設け、この操作スイッチを操作することによりカテゴリを指定する構成としても良い。また、上記実施例では、入力音声で表わされる語彙のカテゴリを指定する処理と、入力音声で表わされる語彙の属性情報を指定する処理を併せて実行するように構成したが、これに代えて、いずれか一方の処理だけを実行するように構成することも好ましい。そして、いずれか一方の処理だけを実行する構成であっても、いずれの処理も実行しない構成に比べれば、認識率が向上する。

【0034】また、上記実施例では、図1のステップS70において、音声認識を再び実行する場合に、音声認識を最初からやり直すか否かを問い合わせ、これに対して、「はい」、「いいえ」という音声で応答したり、「YES」キー、「NO」キーを操作して応答したりしたが、これに代えて、次の通り応答しても良い。即ち、ステップS70において、ユーザーにより、発声する語彙（言葉）の属性情報（属性データ）を指定する操作があったときには、音声認識を最初からやり直さないという応答

があったとして、ステップS70にてNOへ進ませるようにも良い。換言すると、ステップS70において、ユーザーにより音声を入力する操作が行われたときには、音声認識を最初からやり直すという応答があったとして、ステップS70にてYESへ進ませるようにしても良い。

【0035】一方、認識辞書の構成を変更することにより、音声認識の認識率を向上させるように構成しても良い。具体的には、認識辞書内の語彙を、使用頻度によって複数の語彙グループ（例えば高頻度認識辞書と低頻度認識辞書の2つのグループ）に分けると共に、これら複数の語彙グループに使用頻度が高い順に優先順位を付けるように構成し、そして、入力音声の認識候補として複数の語彙が抽出されたときに、優先順位が高い語彙グループに属する認識候補を認識結果として出力するように構成した。この構成によれば、優先順位が高い語彙グループに属する言葉は使用頻度が高いから、それだけ、音声認識の認識率が向上する。尚、認識辞書内の語彙を、使用頻度によって3つ以上の語彙グループに分けても良いことは勿論である。

【0036】また、認識辞書を上記したような使用頻度に対応した語彙グループに分けて音声認識を実行する構成を、前述した実施例に組み込むように構成しても良い。具体的には、前述した実施例で使用する認識辞書として、使用頻度に対応した語彙グループに分けた構成の認識辞書を用いれば良い。このように構成すれば、音声認識の認識率がより一層向上する。

【0037】尚、上記実施例では、本発明の音声認識装置2をカーナビゲーションシステム1に適用したが、これに限られるものではなく、例えば車載用空調装置（いわゆるカーエアコン）、カーオーディオ機器、室内用空調装置、携帯型ナビゲーション装置等に適用しても良い。また、パワーウィンドウの開閉操作の指令や、ミラーの反射角度の設定操作の指令等を音声で行うように構成する場合が考えられ、このような構成に本発明の音声認識装置を適用しても良い。

【図面の簡単な説明】

【図1】本発明の一実施例を示すフローチャート

【図2】カーナビゲーションシステムのブロック図

【図3】ナビゲーション装置のブロック図

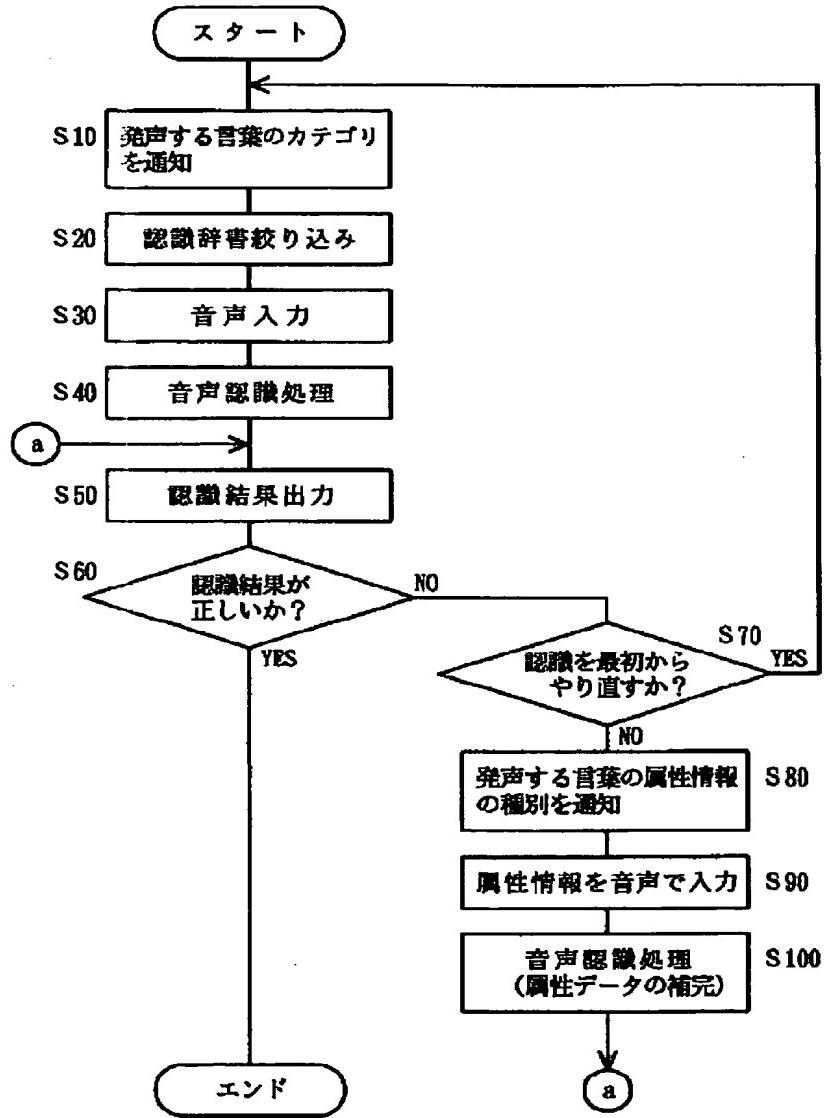
【図4】音声認識装置のブロック図

【図5】音声認識部のブロック図

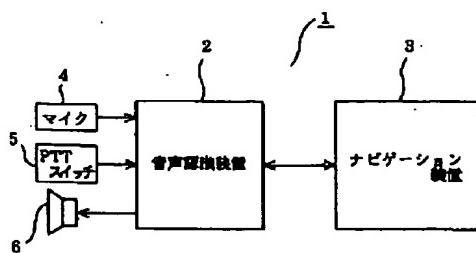
【符号の説明】

1はカーナビゲーションシステム、2は音声認識装置、3はナビゲーション装置、4はマイク、5はPTTスイッチ（カテゴリ指定手段）、7は制御回路、10は操作スイッチ群、14は制御部、15は音声入力部、16は音声認識部（音声認識手段）、17は音声合成部、18は照合部、19は辞書部を示す。

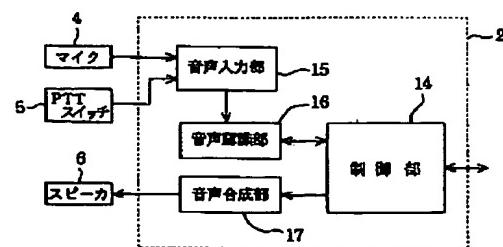
【図1】



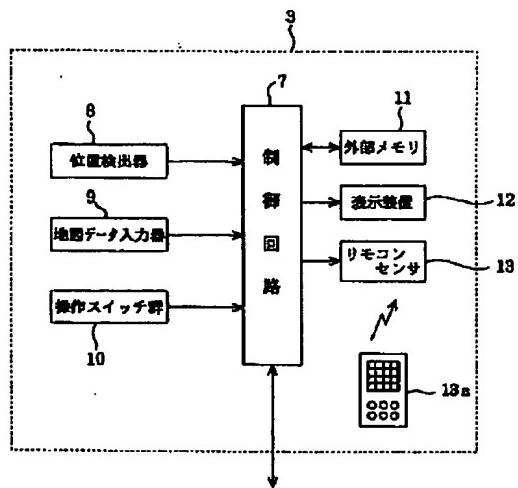
【図2】



【図4】



【図3】



【図5】

